

Decision Making and Aging

Bethesda, Maryland
July 14–15, 2004

WORKSHOP SUMMARY

National Institute on Aging
National Institutes of Health
Department of Health and Human Services

For Administrative Use

Prepared by:
Rose Li and Associates, Inc.
October 2004

**National Institute on Aging
Workshop on Decision Making and Aging
July 14-15, 2004**

Executive Summary

There has been a noticeable trend in the growing complexity of decisions being faced in old age, e.g., pension and benefit issues, portfolio investment decisions, pharmaceuticals, and health insurance options. These are cognitively burdensome decisions, and “information overload” can overwhelm people’s abilities to absorb, process, and weigh information, which can lead to more risk-averse decisions for the elderly. These trends motivate the need for a greater understanding of cognitive issues, particularly with respect to expectations, preferences, and utility – all of which are critical to decision making.

Within this context, the Behavioral and Social Research (BSR) Program of the National Institute on Aging (NIA) convened a small working group of seven outside experts to share ideas in the area of decision making and aging. Specifically, what is the current state of research in this area? What important lessons from the past must guide future research? Where are the opportunities for research advances? And how can research goals best be realized?

The workshop presentations highlighted the importance of affect and motivation on judgments, probability perception, and decision making. Age differences in affective/experiential and deliberative processes have important theoretical implications for judgment and decision theory and important pragmatic implications for older-adult decision making. In some situations, age-related adaptive processes, including motivated selectivity in the use of deliberative capacity, an increased focus on emotional goals, and greater experience, may moderate age-related declines and actually produce better decisions for older adults. Indeed, there is no overwhelming evidence that decision-making quality is necessarily poorer for older people, compared to their younger counterparts.

There is some evidence of a “positivity effect,” i.e., that chronological age is associated with heightened attention to emotional gratification and the emotional aspects of life, and that attention and memory for positive information are accentuated. Researchers are clearly concerned that decisions are influenced by biased recollections of the past that favor the positive, which can place older people at a disadvantage. If the emotional system is improving across adulthood, while deliberate, cognitive processes are declining in efficiency, it remains an open question whether older adults can and should be encouraged to rely upon intuition in order to improve decision making.

The context of the situation (in other words, the “feel” about the probabilities faced) can strongly influence decisions. For example, utility ratings of health states with functional limitations are much lower for people without functional limitations, as compared to those for people who are actually experiencing functional limitations, particularly in younger populations. Patients with

chronic conditions and other low health states also assign low health states to other people in hypothetical scenarios. Self ratings, therefore, appear to reflect one's own preferences, not necessarily just one's state of own health. In some contexts, one's preferences can matter more than his or her health state. Implicit in these findings is the importance of developing tools to help people make better decisions.

In addition to affect and motivation, there are other potential biases that compromise probability assessments, decision making, and information access. "Anchoring effects" appear to be large for people who are highly uncertain, but small for those who are more certain of an answer. There also seems to be an intrinsic bias among respondents in using exemplars in order to create probabilities. In addition, the actual dynamic between the interviewer and the respondent can lead to an acquiescence bias, in which more agreeable respondents seek to please the interviewer by answering positively to any question.

The workshop discussions surrounding perceptual biases underscore the potential contributions that behavioral scientists can make to improve the measurement of utility and cognitive functioning in large population surveys. Several underlying themes permeated the discussion, including the following:

The need for greater cross-disciplinary understanding

The meeting represented an effort to bring behavioral scientists and economists together to help with the survey methodology that many economists use, in order to bring to light differing perspectives on methodology, and foster cross-disciplinary work between these formerly disparate groups. There is interest in facilitating this kind of interaction in the field more broadly, and there are implications for training, as well as for translational research. The behavior-economics initiative is one venue that potentially could strengthen the influence of BSR-type research, with respect to policy and program implementation.

The need to identify common problems of interest

In this meeting, psychological research was characterized as less pragmatic in general than economic research. It was observed that nothing akin to the National Council of Economic Advisers exists to reflect the state of knowledge in psychology or the behavioral sciences. Nonetheless, the field of economics is not devoid of behavioral elements. Without needing to agree on "theology," a particularly promising way for psychologists and economists to interact is to identify an experiment or task that they agree is valuable. Support for such joint research, which does not fit neatly into the existing portfolios, should be encouraged.

The need for better models

To date, the models of age-related changes in factors that affect the perception of affective information (e.g., time perspective, culture, life circumstances, health, time of day, changing physiology, late gene expression) are not as strong as is the emerging interest in them – nor are these models as strong as the surety with which they are presented. It is important to develop research initiatives that would strengthen these models. The potential for BSR guidance in developing and applying the research in this area is profound. The behavioral link to economics and medical decision making would seem to have great potential for translational research; there may be others.

The need for better cognitive data

During this meeting, the Health and Retirement Study (HRS) was discussed, with respect to the potential for adding the HRS survey items that are related to decision making. In this way, the HRS itself may be considered a performance instrument, providing insights into how people think. The major cognitive initiative within the HRS began in approximately May, 2003, and the ideas from that meeting quickly made their way into an approach for taking a smaller number of items that represent latent constructs (e.g., fluid intelligence as a proxy for executive function) and mapping them onto a much larger number of items in a larger validating study that would be separate from the survey. This would be a good way to use existing datasets, where constructs have been well-defined in order to determine the validity of adding smaller numbers of questions to the surveys. Although most of the content in the HRS is fixed, there are many ongoing opportunities to add experiments which would require only a few minutes of questions (or interim mail surveys) directed toward a random subset of the HRS population, or included in HRS-comparable studies in other countries.

One way to improve a substantive understanding of behavior stems from the ability to follow people over time, to look at what happens to their beliefs and their medical conditions. The addition of subjective probabilities has produced an interesting body of data regarding subjective beliefs and the evolution of those beliefs. The success of measurements varies considerably, and there is enormous room for improvement. However, the beauty of surveys like the HRS is their juxtaposition of perceptual with behavioral observations, which provide opportunities to validate measures that are critical to the understanding of decision making.

###

**National Institute on Aging
Workshop on Decision Making and Aging
July 14-15, 2004**

SUMMARY OF PROCEEDINGS¹

Introduction

To date, the strength of the models regarding age-related changes in the factors that influence the perception of affective information (e.g., time perspective, culture, life circumstance, health, time of day, changing physiology, late gene expression) has not kept pace with the strength of the emerging interest in these areas. It is important to develop research initiatives to strengthen these models, particularly since the behavioral link to economic and medical decision making holds great potential for translational research that would improve the quality of life for older adults. Toward this end, the National Institute on Aging (NIA)'s Behavioral and Social Research (BSR) Program convened a working group of seven outside experts, to share ideas in the area of decision making and aging. Attendees were drawn from the fields of psychology (Dr. Laura Carstensen, Stanford University; Dr. Ellen Peters, University of Oregon), medicine (Dr. Alan Garber, Stanford University; Dr. Peter Ubel, University of Michigan), and economics (Dr. Charles Manski, Northwestern University; Dr. Daniel McFadden, University of California, Berkeley; Dr. Robert Willis, University of Michigan; and Dr. Alan Garber). The workshop represented an effort to explore how behavioral scientists might help improve the survey methodology that many economists use, in order to (a) bring to light differing perspectives on methodology, and (b) foster cross-disciplinary work between the groups. The exploratory workshop participants were asked to consider the current state of research in the area of decision making and aging, as well as important lessons from the past that should guide future research, the opportunities for research advances, and how research goals might best be realized.

The workshop format was informally structured in order to allow each invited expert to present his or her own viewpoint, followed immediately by a discussion of each perspective. The workshop also allowed ample time for discussion of common themes and future research directions. Invited experts were asked to provide a few recommended readings for those who might be newly acquainted to their area of research; these readings are included as Appendix A.

Dr. Jeffery Elias began the workshop with observations about human decision making and its limitations. Although some decision making is "hard wired," humans nevertheless can be so focused on a particular outcome or salient idea that they fail to consider all possible outcomes. At the same time, humans rely upon multiple sensory inputs and can function as highly efficient and active information processing systems. Thus, people's first inclination is to "cheat" or take

¹ The findings and views reported in this document reflect both individual and collective opinions of the workshop participants and not necessarily those of the National Institute on Aging, National Institutes of Health, or the U.S. Department of Health and Human Services.

shortcuts, in order to solve a problem as quickly as possible. There are a number of strategies humans use to increase “payoff,” focus efforts, and lessen the workload.

- We store information in multiple systems, tagging information for affective components.² A very exciting area for research is how this tagging occurs, whether this is done bottom-up or top-down, i.e., whether we assign salience to a particular piece of information unwittingly or intentionally, and if the latter, whether this ability changes with age.
- We process information differently within different systems, and store information in multiple formats. An intriguing question is whether we actually have control over how information is stored, organized, or retrieved in our brains.
- We use previous knowledge to process new knowledge, and we mark what is new and what is old; this seems to be hard-wired. For example, we pick out patterns – and, in so doing, we can distinguish deviations from a familiar pattern that spell trouble, as may occur when one is driving. We also survive by using knowledge to actively adapt to continual inputs; this in turn creates a new type of information, which is generated by self action, as well as by immediate and delayed outcomes.
- We are adept at processing information within a context that extends from a macroeconomic (e.g., culture) to a microeconomic (e.g., speed of information) context. We perceive the contextual components of information as being organized in a hierarchical fashion, due to the nature of our processing systems. In reality, the interaction of context may not be hierarchical, additive, or linear, but rather nonlinear, non-additive, and dynamic.

There are several limiting factors that contribute to cognitive changes with age, which have an impact upon strategy selection, frequency of use, execution, and monitoring – all of which contribute to the ultimate effectiveness of our strategy repertoire. These limiting factors include declines in sensory function, memory, speed of processing, motor functioning, mobility, and executive function; changes in the process of affective information-processing; changes in the acquisition of new knowledge; and changes in health status. Previous knowledge overcomes many of these limiting factors. However, the acquisition of new knowledge may not necessarily improve decision making in old age. Indeed, the workshop presentations highlighted the importance of affect and motivation upon judgments, probability perception, and decision making, and they underscored the potential contributions that behavioral scientists could make to improve the measurement of utility and cognitive functioning in large population surveys. These presentations and discussions are summarized below.

² “Affect” is comprised of all the positive and negative feelings one has about a given object, option, attribute, or event.

Invited Perspectives

Ellen Peters, University of Oregon

Choosing, Feeling, Thinking: How Do Older Adults Decide?

A combination of theory, empirical data, and real-life phenomena support the existence of two parallel and intersecting modes of information processing: deliberative and experiential. Epstein (1994) characterizes the former as analytical, logical, conscious, slower, and deriving from fairly recent evolutionary history; the latter is affective, intuitive, holistic, faster, and based upon experience. The psychology of judgment generally has viewed decision making as emerging from only the deliberative system. However, decisions clearly are influenced by the experiential system, as well. Consider the buyers who are willing to pay almost 25 percent more when they are in a positive, as opposed to a negative, mood (Peters, Vastfjall & Starmer, in review). The notion that information processing is based upon not only a deliberative process but also (and equally importantly) upon an affective or experiential process has implications for decision making and risk communication, as well as for how information might be presented in order to help people make better decisions.

In complex decisions, we are restrictively rational. According to this “high reason” view of decision making, older adults should become inevitably worse at making choices. Indeed, older adults do perform less well on unfamiliar tasks (e.g., comprehension of tables and charts). Deliberative efficiency also declines with age. Nonetheless, Peters argues, accumulated experience can compensate for these decreases in performance; in addition, emotional focus can be seen as an adaptive process. As people age, affective information may become more important. For example, with age there is an increase in attention to emotional content, and particularly positive information. Older adults also may be better able to learn from affective cues, as compared to younger adults (Hess et al., 1996). Affective markers guide decision making and the perception of information (e.g., Damasio 1994; Loewenstein et al., 2001; Peters & Slovic, 1996, 2000). It is thought that the markers operate covertly, are learned through experience, and improve the efficiency and quality of decision making. Indeed, affect acts as a source of information. Consideration of possible bad outcomes can lead to unpleasant gut feelings or alarms; consideration of possible good outcomes can lead to pleasant feelings, or beacons of incentive.

The way unfamiliar information is presented also can help determine its utility. A variety of methods can be used to make numbers more meaningful, such as ordering information and introducing memory-based aids in order to lower cognitive effort; using narratives and vignettes or tailoring information to foster greater resonance with the experience being described; and using covert or overt affective markers in order to highlight affective meaning. Even subtle changes in the way information is presented can produce overt affective markers. In one study of hospital plans, for instance, the addition of affective categories influenced choices, particularly for participants with slower processing speed (Hibbard & Peters 2003). These results underscore the salience of affect in perceptions of probability and value; they also are consistent with Prospect Theory, which states that options are evaluated in terms of subjective values and likelihoods (Kahneman & Tversky 1979). Emerging research suggests that parameters may need

to be modified for affect-rich information, as strong affect can be insensitive to probability (Hsee & Rottenstreich 2004; Rottenstreich & Hsee 2001). The finding that older adults may focus more upon emotional content can be problematic for informed decision making in health and financial domains.

In general, decision makers appear to value the world around them by using deliberation as well as affect. While these two types of information processing frequently work in concert with one another, there also appears to be a tradeoff between the two. In particular, reliance upon feelings seems to increase with less cognitive efficiency, due to older age or slower processing of information, among the less numerate, as well as in the presence of affective markers. Reliance upon feelings also increases with cognitive load or stress, which can derive from a variety of factors, including the complexity of the task and information, memory demands, time pressure, and poor health.

Peters concluded by proposing the possibility of an expanded theory that integrates both decision making and cognitive aging, and their interactions. Information processing in decision making appears to evolve throughout one's life. Age-related increases in affective processes may lead to wiser decision making, at least in some contexts. Different decision aids may be needed for older and younger adults. Particularly for older adults, information providers cannot present "just the facts." Every method of data presentation influences choice; choices of data presentation therefore need to be well thought-out and defensible. In communicating risks and benefits, treatment options, and other health- or financial-related information, numbers are just numbers – but affect gives any decision meaning and can provide the motivation to take action.

Laura Carstensen, Stanford University ***Socioemotional Selectivity Theory***

Motivational changes across the lifespan can have profound implications for decision making. According to the Socioemotional Selectivity Theory, humans are uniquely able to monitor time – including lifetime – and they are able to do so at both conscious and subconscious levels. Goals always are set in temporal contexts. Because chronological age is associated with time left in life, it is reasonable that goals should change across the life span (Carstensen 1993; Carstensen et al., 1999). Thus, when time is perceived as expansive (as it typically is in youth), people are interested in meeting new people, because it is unclear what kind of information is needed; on the other hand, when time is perceived as being constrained in some way (i.e., as it is for AIDS patients), people tend to stop acquiring new information and instead focus upon regulating emotional states.

A number of postulates stem from this theory: (1) the perception of time influences goals; (2) goals influence cognitive processing; and (3) pursuing emotional goals benefits the regulation of emotions. Research suggests it is likely that emotional functioning improves into late adulthood and is at least maintained in old age. Indeed, older adults have the lowest rates of virtually all psychological disorders, as compared to younger adults. Older people report that they can control their emotions more effectively, experience more mixed emotions, and feel negative emotions

for shorter durations. Even in the midst of cognitive decline, there appears to be an improvement in the emotional system, at least in terms of regulating emotion.

While emotion and cognition cannot be clearly delineated, emotion is inherent in all cognitive functions. The observation that “categories exist because some objects and events matter and others [do] not” (Robert Zajonc 1997) describes an inherently emotional process. This may be particularly important for the psychology of aging. To the extent the categories that matter change with age, we may see systematic changes in cognitive processing as a function of motivational factors.

A number of empirical studies demonstrate how emotions may matter. In one study, Carstensen and her colleagues presented two sets of advertisements that were identical in terms of their graphic design, but differed in terms of the promises they made – i.e., one set focused upon the future and the other was affect-rich. In the set showing an image of a watch, the caption in the first ad read: “Success is within reach. Don’t let time pass you by,” whereas the caption in the second read: “Take time for the ones you love. Don’t let time pass you by.” The other set of ads showed a collage of photographs, one with the caption: “Capture the unexplored world,” and the other reading: “Capture those special moments.” Whereas the framing of the ad made no difference for the younger adults, older people clearly preferred the more emotional ads. They also were able to remember the emotional slogans more easily than the knowledge-related slogans and were more likely to remember the products associated with emotional slogans (Fung & Carstensen 2003).

In a subsequent experiment, the researchers asked old and young subjects to imagine that they had just received a phone call from their physician virtually ensuring them that they will live another 20 years in good health. Based upon this assumption, the groups were again shown the advertisements; when subject to the new, time-expanded condition, older subjects no longer significantly preferred the emotional ads over the information-relevant ads.

If people are attending more to regulating their emotional states, it is possible that all emotional information is of relevance. In fact, older people show a higher preference for positive versus negative emotional messages. Focusing on the positive can make one feel better in the moment. Carstensen believes there is a “positivity effect,” i.e. that chronological age is associated with paying heightened attention to emotional gratification and the emotional aspects of life, and that attention to and memory for positive information are accentuated. She noted that she does not use the term positivity “bias,” because the effect is driven as much by older adults’ preference for positive information as it is by younger people’s preference for negative information. Accuracy is found to be the same when comparing older and younger adults.

To test the positivity effect, Carstensen and her colleagues presented subjects with stimuli that were positive (bunnies), negative (cockroach on pizza), and neutral (empty bowl), and then asked them through recall and memory tests to recollect the images represented. Older people were more than twice as likely to remember positive images over negative images and they remembered even fewer of the neutral images. Young people remembered more of the images, but were slightly more likely to remember a negative image over a positive one. They also were less likely to remember a neutral image. Middle aged people remembered more of the positive

images over any other age group and remembered slightly less of the negative images than the younger age group. Middle aged people in the study also remembered less of the neutral images than the young (Charles, Mather & Carstensen 2003).

Carstensen pointed out that in most cognitive aging studies on memory, neutral images are compared to avoid confounding from valence. The recall pattern clearly differs when valence is introduced, and these insights are enormously important for communicating effectively with older adults.

A replication of this study – with the addition of brain scans to look at brain activation while the images were being viewed – confirmed again that older adults had a significantly greater recall of images that elicited positive emotions, as compared to recall of the neutral and negative images (Mather, Canli, et al., 2004). The most common response among older people regarding how to deal with negative emotional issues is to “just not think about it.” The brain scans reinforced this view – i.e., that negative emotions simply do not register in older adults, as compared to younger adults (Ibid.).

A recent study examined the positivity effect in autobiographical memory, by going back to an order of nuns who had completed an extensive questionnaire 15 years earlier. In 1987, the School Sisters of St. Francis completed a 26-page questionnaire about their physical, social, and emotional lives. In 2002, they were re-contacted and completed the questionnaire again – “as you completed it in 1987.” By the time of the second questionnaire, the ages of the nuns ranged from 52 to over 100. The researchers used a quasi-experimental design to see if they could eliminate the positivity effect by having the participants focus on how they felt as they completed the survey, as well as focusing upon being accurate. They embedded several prompts throughout the questionnaire, reminding them to be as accurate as possible and to pay attention to their feelings. The younger nuns responded more negatively the second time they completed the questionnaire, while the older nuns answered more positively the second time (Kennedy, Mather & Carstensen 2004).

The findings from these studies have a number of implications for decision making. There is suggestive evidence that older people are deferring decisions in order to avoid emotionally-charged decisions. Researchers are clearly concerned that decisions are influenced by biased recollections of the past, favoring the positive – which can place older people at a disadvantage. If options are presented, older people tend to look longer at positive, rather than negative, options. If the emotional system is improving across adulthood, and deliberate cognitive processes are declining in efficiency, it remains an open question whether older adults can and should be encouraged to rely upon intuition to improve their decision making skills.

Peter Ubel, University of Michigan
Influence of Affect on Probability Perception and Decision Making

Ubel presented three scenarios to exemplify the influence of affect on probability perception and decision making.

In scenario one, subjects were asked to select between two surgery options with the same probability of cure, but with different probabilities of death. Surgery 2 is associated with a lower probability of death but a 4 percent chance of a complication with the cure.

Option:	Surgery 1	Surgery 2
Probability of cure	80%	80%
Probability of dying from colon cancer	20%	16%
Probability of complications with cure	None	4% (1% intermittent diarrhea, 1% wound infection that takes a year to heal, 1% colonoscopy, 1% chronic constipation)

In general, people tended to underestimate how well they will adapt to health problems, for example to severe spinal cord injury. Many chose Surgery 1, suggesting that they do not understand the numbers. Approximately 92 percent favored Surgery 2, choosing complications over death; this suggests that approximately 8 percent of the population would prefer death to complications. A 4 percentage point difference in the probability of mortality is actually quite substantial. Somewhat surprisingly, approximately 40 percent of a group of physicians chose the first option. That people would prefer death over complications is probably due to affect. When respondents are confronted with their apparent inconsistency, results are mixed. To some, a 4 percentage point difference is perceived as negligible.

Although Ubel and colleagues have not examined patterns by age, workshop participants speculated that there would be differences in response by age, i.e. adolescents would probably tend to select the first option; older people who have already experienced some functional decrements would more likely select the second option. The willingness to go on with various decrements changes with age, because people’s goal posts change. Another element is how differently medicine is practiced geographically (Wennberg & Cooper 1999), suggesting huge differences in procedures, which must reside within the medical system framework, and supersedes the effect from many individual decisions.

Scenario 2 involves the perception of one’s average lifetime risk of being diagnosed with breast cancer. When people were asked what they believed to be the percentage of women being diagnosed with breast cancer in their lifetime, the average response was 41 percent, with several estimates of 50 percent. When women learn that the actual figure is 13 percent, they are less interested in getting a mammogram. Conversely, when a group of women were not asked to guess but were told the risk is 13 percent, they stated that the number was about what they expected and that it made them anxious, not relieved. The number “felt” different, depending upon whether the subject was asked to guess first.

In the third scenario, subjects were told that they would have a 6 percent chance of developing cancer in the next 5 years. There is a pill (e.g., tamoxifen) that cuts the chance of developing cancer down to 3 percent, but it has side effects such as hot flashes. The average person has a .5% risk of getting a different cancer that this pill will cause. They were then asked if this were a pill they would want to take. The subjects were then given an additional piece of information: the

average person their age does not have a 6 percent 5-year risk; the average person has a .5 percent 5-year risk. This information should not matter, but seemingly does, and can be alarming to some. The feel of the whole decision is therefore dependent upon other information provided and the context of the situation.

Ubel echoed the earlier presentations, focusing upon the role of affect in decisions and judgments. These three scenarios show that the way a person makes decisions depends upon whether comparative risk information is provided. Perception of probabilities, as well as how people “feel” about the probabilities they face, can strongly influence decision making.

It appears that people’s actual judgments, decisions, and actions do not necessarily match up to the norms set by rational choice theories. People do not make good probability estimates. There are all kinds of biases and anchoring effects, which compromise probability assessments, decision making, and information access. People take shortcuts when trying to combine probabilities and utilities; they may look just at the dominant attributes, instead of a host of separate traits. Ubel cautioned that psychologists’ tendency to study anomalies is insightful, but it cannot be generalized to the larger population. Economists, on the other hand, do not focus as much upon anomalies, but see them as a skewed subset of an underlying mechanism. Manski offered that a standard dynamic programming model by economists with finite horizon would show optimal decision changes with a shrinking timeline; there is less of a time period in which to reap the rewards of any investment in human capital.

New research on the role that affect plays in the way people make decisions or view probabilities includes studies on mis-predicting utilities or the emotional reaction to events; “mis-feeling” the probabilities; and feeling that emotion is influencing decision making. Emotion can affect and even overwhelm people’s abilities to absorb, process, and weigh information. Relatively subtle details can make a difference in probability perceptions and how decisions are made. When outcomes have an affective component, probabilities seem to matter less. When making decisions for others, people tend to make better decisions, because they are more dispassionate.

Garber noted that, from a physician’s perspective, the more complex the decision, the more serious the disease, and the greater the negative emotion associated with the decision, the more the patient wants to defer to the physician’s recommendation. Although there can be comfort in benevolent paternalism, the flip side is the potential for manipulation. It is not clear whether people are more likely to delegate decision making to a trustworthy agent when they recognize that they are making inconsistent choices. Another way to phrase the question is to ask whether a patient would prefer a doctor who has a comprehensive grasp of data or one who relies upon gut instinct. Patients do not want cold, sterile data; they want cold facts to be presented in a warm, fuzzy manner, which underscores the importance of affect. It was noted, however, that the psychology literature is filled with anomalies. Manski observed that people do not always make important decisions on the spot. Some patients walk out of their first visit already sure of how to proceed, while others take time for reflection and active information-searching.

The patient population clearly is not homogeneous. Some patients will question the physician and participate actively in decisions about their own care. Although its use is not as prevalent among older people, the Internet also has changed the physician-patient relationship, because it

empowers patients by facilitating information access and self-diagnosis, serving in some cases as a quasi-agent, which leads to some concern, since not all information on the Internet or in published articles is accurate, and some may be biased.

Studies indicating that older people more often delegate decisions about their health to their physician could be documenting a cohort, rather than an age effect. Although a number of studies show that older people make decisions differently than young people do, there is no overwhelming evidence that decision-making quality is necessarily poorer among older adults than it is for younger people.

Carstensen and Peter's work suggests that people may be viewed as Bayesian, in which their prior knowledge and accumulated experience influence the way in which they determine probabilities. It is well known, however, that determining probabilities can be complicated and tricky in complex situations; indeed, the use of emotion may be a flawed proxy for this type of Bayesian calculation. The technical fix (e.g., formal analyses) might be more effective, but in the absence of that, the proxy works comparatively well. The advantage of using emotion is that it can aid decision making from a variety of standpoints. Emotional responses are informed by experience: affect and cognition often evolve in tandem. Given this, it may be possible to determine an optimal mix of emotion with deliberation.

It was observed that people have different cognitive styles: some are "lumpers," some are "levelers," and others are detail-oriented and even obsessive. Extreme emotions (like high levels of anxiety, or the shock of being diagnosed with a major disease) can undermine the quality of one's decisions and/or make decision making more difficult. It is conceivable that there are several different modes or parts of the brain involved in decision making, some of which may be automatic or unconscious processes.

Alan Garber, Stanford University
Measuring Preferences and Disutility of Functional Limitations

When the process of decision making is approached from the perspective of cost-effectiveness and decision analysis, an intriguing question arises, regarding the value one gives to improvements in the quality of life. As an extreme example, one of the most contentious policy issues Garber dealt with early in his career involved Ceredase, a drug studied by NIH scientists and manufactured by a private company for its exclusive use in treating a rare genetic disease called Gaucher (pronounced GO-shay) disease. Given that Ceredase costs \$350,000 to \$500,000 per year, the drug is not cost-effective.

Since then, Garber has become more interested in quality-of-life issues and the elderly, motivated in part by a randomized trial of the Geriatric Evaluation and Management Units (GEMU). The GEMU program postulates that elderly people with multiple impairments might benefit from the targeted forms of multi-disciplinary care that are embedded in GEMUs. A randomized trial has shown a large decrease in mortality (although this result has not been replicated in subsequent studies), as well as improved functioning among those who received this intervention of multidisciplinary care.

In order to study the value to older people of improved functioning, Garber and his colleagues developed an elaborate system of evaluating impairments in daily living. One of the first challenges was identifying which of the many activities of daily living (ADL) the utility assessments should apply to, and how to select the appropriate scenarios involving people with ADL limitations. The researchers then developed a process to ensure that everyone involved in the preference-assessment exercise rated each of the relevant ADLs, as well as a random combination of a limited number of other ADLs. Data from the National Long Term Care Survey were used to estimate the frequency of different combinations of the ADL limitations, so that the researchers could give appropriate weight to the most frequently cited ADLs. Over the years, Garber and his colleagues have developed software to assist with measuring preferences. It shows patients detailed, multi-media scenarios that depict what it is like to live with a specific disease and how it would limit one's activities. Patients then are given standard tradeoff scenarios, and researchers assess their preferences. This exercise has led to a number of interesting findings.

Garber reported that, in his studies, people routinely gave inconsistent responses. For example, one would expect that respondents who assigned high disutility values to each of two ADL limitations would assign a higher disutility value to the combination of these two ADL limitations, but this was not always the case. If such inconsistencies are pointed out, however, respondents seemed ready to change their responses.

Other ongoing research that is part of an NIA-funded study with a demographically representative population in California (Kaiser) has led to a number of interesting findings that relate to the earlier discussion:

1. Disutility ratings vary according to a person's age and his or her experience with functional limitation. People without functional limitations gave much lower utility ratings to health states with functional limitations, as compared to those people who actually were experiencing functional limitations; this was particularly the case among younger populations. Although utility ratings were consistently much lower among people who had not experienced a condition, this pattern appeared to be reversed during old age. In a life care facility in Palo Alto, California, a study was conducted with senior housing residents who had extensive exposure to others (e.g., spouses, friends, other family members) with functional limitations, Garber reported that, as people grow very old, the disutility of having a functional status limitation increases (i.e., utility declines), which is consistent with Manski's dynamic programming model. Based upon anecdotal interview data, it seems that people believe they will not recover and are on a downward trajectory. In this regard, younger people (including quadriplegics) differ from the very old; they are more optimistic and have the attitude that they will be able to make adaptations in order to take care of themselves. Garber hopes that this Kaiser sample population will enable him to ascertain whether or not his findings hold up with a wider age range.
2. Preferences – not just health states – drive utility ratings. There is a distinction, however, between quality-of-life and utility rating. Garber is seeking an objective assessment of health impairments and health states. Patients with chronic conditions

and other low health states also assign low health states to other people in hypothetical scenarios. It is clear, then, that self ratings of health reflect one's own preferences, and not just one's own state of health. It should be noted that there is an overlap between healthy and supposedly unhealthy persons; preferences can matter a great deal, and in some cases may be as important as a patient's actual health state.

The software that Garber has been developing to help measure ADL preferences is at a fairly advanced state. It reflects a combination of disease and health state descriptions as well as preference assessments, and it permits researchers to check for consistency. Garber reported that he is prepared to share this software, if it would be useful to others. He also noted that other large-scale utility assessment exercises focus upon very specific conditions, such as wearing glasses or having headaches. In contrast, his research has focused upon conditions that are associated with long term care, and this can be of value in the field of aging research. Garber explained that he concentrates upon patients with severe limitations, because it would be impossible to capture the whole range of ADL limitations.

In terms of assessment techniques, Garber remarked that respondents hate dealing with "standard gamble" approaches (i.e., giving people a lottery choice to find out what value they place on improving their health state). When timed, it took respondents longer to answer questions that involved using a gamble-based scale, as compared to questions that used the rating scale. The numbers generated (e.g., a value of .5 on a scale of 0 to 1 = you would give up 2 years of life in this health state in order to have 1 year of life with unimpaired health) were so far off the rating scale as to be implausible. (In this situation, if subjects were told that their numbers were way off, they reconsidered their choices.) Even though studies showed that respondents disliked the standard gamble-based questions, their responses were quite reasonable once the concept was explained, whereupon the respondents tended to be more comfortable with the questions. In general, one might say that the more painful, difficult, and time-consuming the assessment tool is, the more likely it is to yield useful the information.

Discussion ensued about so-called anchoring effects, which might be a concern if one were to consider adapting some of these methods to the Health and Retirement Study (HRS). Garber has invested a great deal of effort in developing his question methods. The choices provided allow the respondent to enter a number; following this, a ping-pong technique is employed to enable respondents converge on an answer. Another method is to give patients obvious choices (e.g., "Would you accept a risk of death of .3 in order to be free of glasses?") — then, if they do not provide the expected answer, return to the survey and educate the respondents, since their answer would indicate that they may not have understood the exercise.

Instead of asking questions devoid of any context, Garber has tried embedding choices in vignette-like situations, in which more information is provided about individuals and their circumstances, diseases, and preferences. Earlier work also looked at providing more as opposed to less information – i.e., supplying information about a person's children and which specific activities they can or cannot do, in order to provide a richer context for the study. Findings from these studies indicate slightly less variance in the answers to these questions, but the provision of such detail did not appear to make a substantial difference.

In general, utilities are not super-additive. That is, if one ADL limitation were given a utility of .4 and another .5, the combination of the two limitations typically would not drop to .1, although one might reasonably expect it to be less than .4. However, a number of people preferred death to having a large number of combined ADL limitations. Even among consistent responders, the utility rating did not go down to .1, leading researchers to conclude that the utilities are not completely additive. Garber considered the responses consistent, so long as the combination of ADLs did not result in a utility rating that was higher than the worst of the two.

It is sometimes difficult to separate learning from affective forecasting – that is, predicting how one is going to feel, which is greatly influenced by learning. Defining learning as the capacity to change one's response based upon new information, we might expect older adults to improve in the area of affective forecasting. However, it is not clear whether learning enables people to update their ability to estimate duration or intensity. Sometimes interventions were necessary in order to remind people about an actual experience. It also was observed that expectations do not always follow learning (e.g., alcohol myopia – a condition for which the weight of the evidence does not at all influence behavior). Studies appear to indicate that, at times, people's expectations get in the way of what actually occurs.

Implicit in Garber's presentation is the importance of developing tools to help people make better decisions. The premise is that physicians can make better recommendations to their patients if they know something about the patients' preferences, as well as the relevant probabilities, based upon particular medical conditions. Since some biases (e.g., context, framing effects, and/or the questions being asked) influence responses, Garber has striven to create an environment that removes as many sources of bias as possible.

Robert Willis, University of Michigan ***Embedding Experiments in Household Surveys***

Decision making has become increasingly important in economics research. Experimental economists typically rely upon opportunistic samples of relatively small populations; therefore, in these types of experiments, little is known about the subjects' history or later outcomes. The focus thus is upon the main effects, not upon measures of heterogeneity within the population. Many ideas which have been investigated by experimentalists could be tested in a broader setting in order to better understand their policy implications – i.e., if the experimental economy ideas were implemented in a particular context, with a richer set of information from people's life histories and with future follow-ups.

On a personal note, Willis described himself as an economist with a longstanding interest in decision making, beginning with his research on the economic determinants of fertility behavior (decisions with very deep implications for prospective parents), as well as his research on the economic determinants of investments in education (that is, in terms of considering both the costs and benefits). As the Principal Investigator on the HRS, Willis became increasingly interested in the available literature on cognition. He observed that many of the issues under discussion occur in the administration of surveys themselves; indeed, large-scale surveys offer fertile ground for psychologists, because they can capture the way people respond in various

situations. Thus, surveys themselves may be viewed as performance instruments, providing insights into how people think. Performance tasks in surveys also present unusual data-collecting opportunities for economists. This has led to a greater interest by economists in subjective probabilities.

Willis provided some background on the HRS in order to give a sense of the various options, promises, and obstacles for embedding experiments in household surveys. He reviewed the core content areas of the HRS, its evolution, and its steady state design. The original HRS cohort (N = 12,654) was first interviewed in 1992 and included persons age 51-61 in 1992 (born 1931-41), as well as spouses of age-eligible persons, with a longitudinal follow-up every 2 years. In 1998, a redesign merged the HRS and AHEAD into a single study. In 2004, the HRS was awarded additional funds from the Social Security Administration (SSA), to conduct most interviews in person, which made the inclusion of performance measures possible. Today, cognitive issues and cognitive measures are playing an increasingly larger role in the study.

One important supplement, which began in 2001, is an in-home study of dementia. The Aging, Demographics, and Memory Study (ADAMS) uses a stratified random sample based upon survey measures of the risk of dementia involving three groups: people with high probability; those for whom dementia is less certain; and those who are apparently normal. The dementia assessment is conducted through case conference by a research group at Duke University that is familiar with these types of assessments. The sample population in ADAMS is very different than a typical clinic population, being usually eight or ten years younger.

Other components of the HRS include a number of mail surveys on (economic) consumption and time-use. Another component seeks to develop an independent Internet sample constituting approximately one quarter of the sample. Willis also noted the linkages to Social Security, health insurance, Medicare, and employer pension plan information, as well as the National Death Index.

Motivations for conducting experiments in the HRS include finding ways to improve data quality and increase response rates. One way to improve a substantive understanding of behavior stems from the HRS' ability to follow people from wave to wave, and look at what happens to their beliefs and their medical conditions. The addition of subjective probabilities has produced an interesting body of data on subjective beliefs and the evolution of those beliefs. Attempts to measure risk tolerance (Barsky, et al., 1997) and time preference (scattered measures) correspond with standard economic models of decision making. The success of measurements varies considerably, and there is enormous room for improvement. In principle, it is possible to think innovatively, even though most of the HRS' content is fixed. Opportunities exist to add experiments that would require only a few minutes of additional questions for a random subset of the HRS population or on HRS-comparable studies in other countries. In addition, self-administered questionnaires might be used as a way to check for consistency. In trying to improve subjective probability measurements on the Internet, it also may be beneficial to use graphical tools. Developing a test bed sample has been suggested for Internet interviewing: for example, using the HRS sample to test for mode, selectivity, and/or response rate effects in order to guide further developments of an Internet-based mode for HRS data collection. This method might be appealing, since there is a high fixed cost in establishing an interviewer and respondent

relationship. On the Internet, one can conduct high-frequency interviews more easily. Establishing a panel to assist with high-frequency interviewing also would shorten feedback time between the experiment and results.

One of the major goals of the HRS is to obtain high quality-measures of income and wealth. Wealth is cognitively difficult to measure. For example, the difficulty of measuring net worth prompted survey designers to seek better answers by breaking the question into parts: Do you own your house? How much is it worth today? Do you have a mortgage? How much is that? Do you have a checking account, stocks/bonds, and/or other assets? Although precision is likely to be enhanced by asking about the components of wealth or income, a respondent may not know or want to share specific pieces of information. Ultimately, economists prefer to develop a measure for total wealth or income, not so much the components. Missing values pose a major challenge, especially in the presence of the enormous heterogeneity of the respondent population, which makes extrapolations to estimate missing values difficult.

Development of the “unfolding brackets” technique has led to a substantial reduction in non-response. However, after the first waves, the HRS investigators were alerted to research papers which suggested that unfolding brackets are vulnerable to anchoring effects (Locander & Burton 1976; Jacowitz & Kahneman 1995). As a result, the HRS team began a series of experiments to determine the extent of the problem and to develop ways to address it (Hurd, 1999).

There is evidence that people interpret the survey as providing information. Anchoring effect appears to be large for people who are highly uncertain and small for those who are more certain. Willis reported that people who answered questions about subjective probabilities that indicate a great deal of uncertainty were giving focal answers (e.g., 0, 50, 100), which had consequences for their net worth reporting.

Investigators have experimented with randomizing the brackets, such that every quantitative question has random entry points so that researchers can study the anchoring effects. Hurd discovered that this created another bias: not just anchoring, but an acquiescence bias, in which more agreeable respondents were more likely to answer positively to any given question.

These findings suggest that the survey itself is actually a transaction between the interviewer and the respondent. There are inextricable connections between the survey indicators of cognitive performance and behavior in the real world. As has been seen in numerous examples, a question asked in different ways elicits different answers. Therefore, survey developers have made a careful effort to alter the sequence of questions in order to minimize the possibility of acquiescence and anchoring bias. As researchers need to apply some interpretation of the data, there is room to develop more behaviorally-based interpretations, and to understand the objects to which people are responding.

There has been a noticeable trend in the U.S. toward increasingly complex decisions which must be faced in old age (e.g., pension and benefit issues, portfolio investment decisions, pharmaceuticals choices, and health insurance options). These are cognitively very burdensome decisions, and “information overload” seems to lead people to make more risk-averse decisions. These trends have not only broadened the scope of the HRS in terms of content coverage, but

also have motivated a realization of the need for greater understanding of cognitive issues in survey administration.

Charles Manski, Northwestern University
Measuring Expectations

Economic models consider expectations about the future to be absolutely critical. Nonetheless, the long-held norm among economists has been to observe people's choices in order to infer preferences, not to collect subjective data about preferences, utility, or expectations. Even in a standard utility model, there could be many different combinations of preferences and expectations behind any particular decision or outcome. Conventional wisdom against collecting expectations data started to break down among economists, beginning in the early 1990s; the HRS was an important (but certainly not the only) survey to develop innovations in this regard. When analyzing economic behavior under uncertainty, it has become accepted to replace assumptions about expectations with self-reported data. "Understanding how persons revise their expectations with receipt of new information often is a prerequisite for credible use of econometric decision models to predict behavior" (Manski, 2004).

Economists generally do not find it appealing to ask for the long verbal explanations that are typically associated with attitudinal research in social psychology. Instead, they prefer to obtain probabilistic assessments. Thus, when eliciting information about respondents' expectations, economists prefer to pose probabilistic questions about well-defined events, rather than verbal questions about vague events.

A number of major economic surveys elicit probabilistic expectations:

- ***Bank of Italy Survey of Household Income and Wealth*** (1989 on): earnings and employment (Guiso, Jappelli & Terlizzese 1992; Guiso, Jappelli & Pistaferri 2002).
- ***Health and Retirement Study*** (1992 on): survival, retirement, bequests, investment returns (Hurd and McGarry 1995, 2002; Hurd 1999; Hurd, Smith & Zissimopoulos 2002; Van der Klaauw & Wolpin 2002).
- ***Survey of Economic Expectations*** (1993-2002): income, employment, social security, investment returns, crime victimization (Dominitz & Manski 1997a, 1999b; Dominitz 1998, 2001; Manski & Straub 2000; Dominitz, Manski & Heinz 2003), versions of which have migrated over into the Survey of Consumer Finance and the HRS.
- ***National Longitudinal Survey of Youth-1997*** (1997 on): schooling, fertility, arrest, crime victimization, mortality (Fischhoff et al., 2000; Dominitz, Manski & Fischhoff 2001; Walker 2001; Lochner 2003).
- ***Michigan Survey of Consumers*** (2002 on): investment returns, income, employment (Dominitz & Manski 2003).

Manski provided an example of a study on mutual fund investment expectations, asking respondents what they think is the percent chance that a \$1,000 investment will increase in value in the year ahead. Dominitz and Manski (2004) show enormous and systematic heterogeneity in "consumer confidence" that the fund investment will increase, despite the fact that the rational expectations model would suggest that everyone should have similar expectations. Instead, there are systematic tendencies by particular demographic groups; older people are less optimistic

about the market than are younger people, and males and those with more years of schooling are more optimistic than females and those with less years of schooling. These findings might be interesting, because different investment choices may depend upon one's expectations. Even if it is not possible to match expectations with behavior from June 2002-May 2003 (after the stock market crash), there is enormous behavioral information in the HRS regarding respondents' portfolio choices. The mean probability of finding an increase was 42 percent. When the same question was asked in 1999-2000, it was 60-65 percent.

Willis noted two design features of the HRS: it not only provides a great deal of information about people's investments, but is also large enough that the field period for one wave extends from February through the end of November. Weekly or monthly samples thus are large enough to detect stock market variations. The HRS began asking subjective probability questions in 2002, and asked them again in 2004, making it possible to obtain responses from the same person at two different points in time: once during the economic decline and again when the economy improved (albeit, with ups and downs). From a theoretical point of view, the rational expectations model would place relatively little weight upon recent experience, due to the long stock market history with basic, expected returns. One of the criticisms of the efficient markets hypothesis is it fails to account for the fact that people give weight to experience.

Manski raised another example involving Social Security expectations. Any economic model of how people decide to allocate their income between consumption and savings depends critically upon their expectations about pensions and Social Security. In principle, one can ascertain a person's subjective probability. Many younger people (aged 30-35) think there is a low likelihood of receiving any benefit from Social Security. There is essentially no age pattern in subjective median of benefits at age 70.

It is important to recognize that uncertainty about future benefits may be due to a number of factors, including the following:

- 1) Uncertainty about future earnings
- 2) Uncertainty about the political outlook
- 3) Ignorance about the formula for calculating benefits, and/or overall ignorance about how the "system" works

Young people appear to worry more about the total collapse of Social Security. (Similar questions could be asked about the Medicare system.) Implicit in these questions is the assumption that respondents will survive to age 70. Another interesting question is whether people ignore Social Security and rely instead upon their own private funding sources. Another aspect is whether they bother to find out more about Social Security. The beauty of the HRS is the simultaneous collection of perceptual and behavioral observations.

Manski's presentation stimulated discussion regarding the extent to which interviewers probe further in the face of unlikely answers, and the extent to which individuals are able to think probabilistically. The HRS has a significant body of expectations questions, and it is very difficult to discern how people actually are forming their expectations; doing so would require much more intensive interviewing. Obtaining a simple measure of expectations is much easier. Manski and Willis are exploring such approaches as using a bracketing technique on 50 percent of the sample, and working in conjunction with psychology laboratories. Manski observed that it

is uncertain whether or not a report of 50-50 really has meaning; the question arises: Is it simply the result of not knowing, but feeling pressured to give a response? Some who respond 50-50 really might mean 40-60. It also is possible that others, who simply do not have well-formed probability beliefs, might think in terms of intervals and are unable to come up with a more refined answer. Much of this might also be due to rounding (i.e., in which a respondent equates 3-4 percent to zero).

In an economist's view, people have expectations about everything – but they may have a great deal of uncertainty. The answer a respondent gives also may depend upon context. In other words, if asked about the probability of something in comparison to other events, a person may provide different answers. One psychologist posited that some adults simply do not think about Social Security. Some psychologists believe that people cannot or do not think probabilistically and need to make their assessments verbally. From this vantage point, most people do not go through life with precise numbers in their head. When asked for a precise answer, people often provide an arbitrary number. Other research has shown that people do not properly update their sense of probabilities when they receive more information, even though their ability to make probability-based assessments may have been quite good initially (Kahneman & Tversky 1974).

According to Willis, the rate of refusal is comparatively low for the HRS probability questions. Non-economists suggested that many people just do not think probabilistically, or doing so is considered to be akin to second-guessing God. One economist countered that nothing in the substantive results suggests that people do not think probabilistically, although it is certainly possible that people cannot articulate their probabilistic thinking. For example, ask a fisherman how likely he is to catch something, and he will have a fairly good idea – which supports the method of correlating responses with actual behavior. Indeed, survival probabilities actually line up quite well with actuarial data (cf. Hurd, et al).

To ascertain whether or not behavior is consistent with utility models, one needs to measure subjective probabilities and behavior in the same dataset. Only the HRS and the NLSY are known to have this capability. Of course, the question remains: when a respondent says 70 percent, does he or she really mean 70 or 75? Is 50-50 truly reflective of the respondent's sense of equivalent chances, or does it reflect that he or she has no clue what the answer should be? It is not clear how refined the responses are. It also is not clear whether the subjective probability numbers are any more imprecise than asset figures.

Economists have a practical motivation: they want to predict people's behaviors (i.e., forecasting consumption decisions) and how they will be affected by policy. Thus, the tools that economists tend to use are expected-utility models. Economists are interested in how people will behave under different scenarios, so they can make comparisons. Economists expressed some frustration with the psychology literature, in which the goal is perceived to be understanding how people think, but not how they behave. Economists now must contend with the possibility that affect influences the reporting of probabilities.

Dr. Richard Suzman suggested that people probably hold several possibly contradictory probabilities simultaneously, which could lead to a refusal to answer, vacillation, conflicted behavior, or answers that are highly dependent upon a question's format or context. Cognitive

style therefore cannot be overlooked. How long one expects to live is a very emotional issue. People may have superstitious beliefs about life expectancy, such as not wanting to “tempt fate” or invite bad luck. There is, consequently, a bias toward maintaining the status quo whenever people feel they do not know the “right” answer. Perhaps econometrics must yield to the psychology lab.

Willis insisted that we need to make this an empirical issue that is conditional upon probabilities, in order to generate mathematical formulae regarding how people should behave. That is why Manski is proposing a “beliefs supplement” to the HRS. It is plausible that people improve their calibration over time and probabilities could change over time as well.

People who are more analytical and not highly subject to biases in judgment tend to seek out information at the time the information is relevant. In a survey context, the information is not, on the whole, realistically relevant. In other words, analytical thinkers tend to investigate information thoroughly when they are in an actual, clinical setting (i.e., when facing the probability of being diagnosed with prostate cancer, they will thoroughly research the pros and cons of having a PSA or any other test). Outside of that context, however, they are less likely to conduct such thorough research.

In light of these findings, Manski argued for the importance of integrating into the survey more personally meaningful events (i.e., those which are generally associated with better answers). If one were to consider life cycle-consumption/savings seriously, then someone who is at least 35 should find the topic of pensions and Social Security to be salient.

Vis a vis the survey questions that are now at the forefront, there is a clear and strong need for effective collaboration between economists and psychologists; these two groups with their disparate viewpoints must delve more deeply into the subject and work together to navigate the process of discovery.

Daniel McFadden, University of California at Berkeley
Choice Behavior in Older Populations

McFadden reported that there is some pattern of systematic bias in subjective probabilities. The older respondents become, the more likely they are to underestimate their probability of survival. One possible theory is that these people are recalling exemplars — i.e., the tendency is to remember the friend who contracted breast cancer, not the many friends who did not. Non-coincidences, then, may be undercounted. Furthermore, in using exemplars to create probabilities, there seem to be an intrinsic bias, which affects the way people collect and retrieve information as their basis for forming probabilities.

Choice behavior in older populations is substantively and methodologically affected by a number of factors, including anomalies in memory, perceptions, beliefs, and judgments. These factors influence choices in economic and health planning and may have a harmful effect upon health and well-being. The methodological effects are associated with cognitive anomalies, which influence responses to the questions found in economic and health surveys, and which may distort analyses.

McFadden is primarily interested in the effect of methodological issues upon surveys. He is researching possibilities for how to embed a person's experiences in surveys, and how to identify the effects. He pointed out that surveys are "structured conversations between strangers," and as such they are subject to most of the same communication problems that can arise in the course of ordinary conversations – including inattention, misunderstanding, strategic motives, posturing, and projection. Surveys require respondents to engage in cognitive tasks that may be misinterpreted or processed incorrectly. In addition, a respondent's retrieval of memories and facts may be incomplete and/or inaccurate – a situation that is analogous to errors made by some people when taking tests.

McFadden's ambition – which some prominent psychologists consider neither possible nor worthwhile – is to find the invariant properties of the entire library of cognitive bias, in order to predict and reverse their effects upon survey results (presuming that they are distorting outcomes). This project may be overly ambitious but certainly interesting.

McFadden then outlined the survey response process (adapted from Tourangeau et al. 2000) that begins with (1) awareness/comprehension/construal; then requires (2) retrieval of facts, exemplars, affective experience; before (3) exercising choice/response/reporting.

Thus, for example, the question: "How much have you spent on food away from home in the past six months?" could be focused upon a number of factors – restaurants only? Fast food? Snacks? Drinks? Food/entertainment packages? Inclusive holidays? Purchases for others? Take-out food consumed at home? Groceries? etc. One might also consider whether there is a significant event or date by which to mark the past six months. One might wonder further about the motivation behind this question – i.e.: Are they trying to determine whether I am a fast food junkie? Overly self-indulgent? Normal? Living a full life? etc.

The processes which respondents use to integrate retrieved information include: judging completeness and accuracy of retrieved memories; making inferences based upon the process of retrieval and to fill in gaps; considering the date, duration, and frequency of judgments; mapping estimates (through decision logic) to choice/response category; and editing one's own responses.

There also are a number of possible response errors. Misreporting of economic facts can arise from each stage of the response process. Survey design can influence errors, perhaps differentially so at various stages of the response process. For example, a question about food eaten in restaurants during last week may be answered more accurately (i.e., by enumeration) than a question about food consumed away from home over the past six months. Finally, known cognitive effects can be influenced by survey design. Another factor is the apparent tendency toward strong anchoring to focal values.

Also related to anchoring, in a recent study using a sample from the AARP, McFadden and Dr. Norbert Schwarz examined order effects. People were asked to state the probability of a person in the United States eventually going into a nursing home, and were given either a higher or lower possible range. The next question was about their personal probability of entering a nursing home. The respondents who were given a higher range for the first question tended to

report higher personal probabilities on the second question. The third question then was: “Have you purchased nursing home insurance?” A significantly higher percentage said they had purchased nursing home insurance, if they had been given the higher range for the first question.

Despite these findings of anchoring effects, McFadden did not conclude that economic surveys are useless; he feels they have too much information of value to be dismissed. Nonetheless, he observed that there appears to be “a minefield” of cognitive effects and ambiguities, which makes the “test bed” concept (i.e. offline experimentation) even more applicable and important, in order to build enough treatments and variation into the survey to mitigate these negative effects.

Discussion

To answer the question about what psychologists need to know before they can enter into any kind of research relationship with economists, we must first prioritize the research questions. The HRS is fundamentally concerned with issues upon which literally hundreds of millions of dollars hinge – e.g., Social Security and Medicare. Many generic problems associated with measurement and psychological issues are involved when asking probability questions. We can learn a great deal about the way people think, then use this information to develop surveys that are better at eliciting consistent and meaningful answers. There are also issues of validation. Subjective health measures can be difficult to validate, although economic measures can be corroborated by external validation data, as can mortality. It would be worthwhile to encourage psychologists to think more about measurement issues and how they can help design better surveys, as well as how to establish a mechanism for disseminating findings. This should be doable, since the HRS has had long experience in working with various expert groups and working within their respective content areas.

An anchoring effect on a single consumption question made the HRS survey designers realize that they simply cannot take shortcuts in the way consumption questions are asked. Because of this, the HRS embarked upon mail surveys and now has the first national estimate of consumption in a longitudinal context. It is extremely important to maintain an ongoing dialogue between survey designers and data users in order to improve measurement. Researchers representing different perspectives need to understand what the other disciplines are doing, what they do and do not believe, and when they believe in the same things but use different terminology. Methodological differences exist, such as the fact that economists are interested in experiments but tend to have a narrower range of experimentation than do psychologists. Surveys offer opportunities for more experimentation. Social scientists are likely to resonate with the idea of embedded surveys, but it will be important to develop parallel processes in order to allow for further experimentation. Innovative ideas should be tried out, and although some might fail, the effort ultimately will help to shape the way the HRS is conducted.

This useful meeting yielded a number of suggestions for moving forward – one of which was to form a working group to more carefully consider the various elements of decision making, how to prioritize questions, and how to develop promising avenues. The end product could be a set of recommendations to the HRS, and presumably to NIA for funding, arguing for improvements to the survey’s subjective measurements. The proposed improvements would be key to a better

understanding of decision making, preferences, and utility – in a manner similar to validation studies (i.e., the ADAMS), which aim to assess the correlation between clinical standards of diagnoses and survey/performance data which indicate cognitive status. If survey questions can predict the clinical dimension for a subset of the survey population, then one might consider the measures as adequate proxies for the entire survey population. One thus might hope to investigate questions over the whole population range, despite the enormous heterogeneity found within the older population.

A planned effort also is underway to assess the extent to which the HRS actually collects data that is useful to psychologists. A cognitive study led by Dr. Jack McArdle will administer (through a telephone interview) the HRS's cognitive items, as well as several basic demography items. As quickly as possible, the same HRS items will be administered to the same subjects in their homes, in order to study mode effects. At the home visit, interviewers also will administer the HRS subjective probability items: a three-hour battery of test items from Woodcock Johnson, as well as other questions. An effort also will be made to re-engineer the HRS cognitive items, so that they address range effects and other potential biases. A baseline sample of approximately 2,500 people would be sufficiently large to study some of the issues discussed at this meeting. Adding the use of functional magnetic resonance imaging (fMRI) might also be considered.

When asked what would happen if economics were introduced into the psychology curriculum, participants speculated that psychologists might pose more relevant questions for public policy or interventions. One economist believed that effective collaborations can occur with some extra background reading and meaningful conversations between psychologists and economists. The policy element of economics, and perhaps some of the statistics (including policy adaptations from anecdotal evidence to statistical analyses) would be most likely to prove useful for psychologists.

Participants were not convinced that giving a group of naïve psychologists a week's course on economics would help them interact more effectively with economists. It is much more important to find questions of common interest, particularly those that are solvable. Policy problems are much more complex, and often are not presented clearly. Economists historically have worked both sides, in order to test basic theories and meet a policy relevance standard. The best approach may be to have representatives from the different fields, who may not necessarily agree on "theology," agree upon the experiment itself. For example, a promising way to interact might be for the parties to mutually agree upon a single, valuable task. The group did not favor the use of training programs, and instead recommended supporting joint research that might not fit neatly into the existing NIH grant portfolios.

Psychological research was characterized as being generally less pragmatic than economic research. One participant observed that there is no "Council of Psychological Advisors," although there is a Council of Economic Advisors. Economists take on questions that seem relevant to policymakers. Clearly, what is viewed as an "interesting" question differs by field. Psychologists sometimes consider applied topics to be less worthy of study. However, there are signs that the incentive structure in the cognitive sciences field is changing. Journals on cognition are increasingly solicitous of research articles that have relevance for policy, practice, and application. Manuscript authors seeking to be published today need to make the case for the

translation of their basic research findings, and they need to demonstrate that their results matter in terms of their practical applications.

McFadden noted that for the last decade Dr. Daniel Kahneman has run a summer program at the University of California, Berkeley that has exposed cohorts of economists to new paradigms of thinking that are not necessarily consistent with how economists traditionally think. Nonetheless economists tend to be unashamed about their admittedly imperialistic tradition, unapologetically seizing the most useful approaches for their own purposes. Economists typically go through a mental checklist about policy effects. Part of this is due to the fact that economic thinkers function in an environment in which they engage in ongoing discussions. Although it may take a non-economist much longer than a week to understand this mindset, an outsider may become an effective collaborator in a relatively short period of time. A much more circumscribed problem may be understanding how economists learn to solve problems. Some meeting participants believed this could be accomplished in a week; others felt that one week is, by definition, too short a period to learn anything effectively.

An alternative approach might be to reach students who can in turn influence their professors. The HRS has been a successful vehicle for shaping various groups, such as the ADAMS group, whose collaborators would not otherwise have interacted with each other. It is important, however, to have repeated interactions, with feedback loops and opportunities for further discussion. One-week courses can be a supplement, but are insufficient in and of themselves.

Another approach might be to form networks. Drs. Willis and McFadden have been planning a “flotilla” project, to form and provide seed funding for several networks, some of which could involve affiliated activities. A number of these activities could be relatively low-cost, while others might be larger projects, such as a joint project about beliefs and decisions involving laboratory experiments, including the use of Internet surveys. Such a flotilla-like mechanism could be a useful way to engage researchers from different fields.

Another suggestion was to present opportunities for vertically-integrated projects which deal with multiple levels of projects in different ways that allow for multi- and inter-disciplinary approaches, and in which one could proceed from basic questions to neuropsychology and neuroeconomics. Such efforts could be embodied either in program projects or in centers. It would be difficult to support such efforts through a regular research (R01) project, unless there were a special Request for Application (RFA) with a special review group. Some small projects could be funded through other sources, or even unfunded: it is not unusual for a small class project to become an important piece of work in the field.

Suzman reminded participants that R03 Small Grants provide up to \$25,000 per module and up to four modules, with an application limit of 10 pages. The R21 application limit is 15 pages and allows requests of up to \$300,000. For centers, there are pilot cores, which are the most flexible forms of funding available. Program projects (P01s) can have a pilot core in which the PI provides overall direction. However, NIA staff cautioned that there are limits to the number and funding amounts of the P01s, and that the specially-convened review group for P01s ensures that it is a particularly rigorous review process. Suzman added that the NIH tradition is to review the application, not the people. He encouraged applicants to consider attaching pilot funds to a

network. Allocating up to \$2,000 for each of a set of commissioned papers is perfectly acceptable in a conference grant that could be part of network activities.

Participants recognized that some psychologists are not at all interested in policy but are very interested in understanding the mind, including the ways in which a stimulus activates a particular region of the brain. It may be threatening to the cognition field to require their research to have immediate policy applications. There are, indeed, many psychologists who are interested in developing basic research within the context of a more practical, policy-oriented domain. Regardless of whether psychologists are or are not interested in policy, however, decades of basic research on how people think has proven to have useful applications in numerous other fields. Clearly, there is good reason to sustain basic research efforts in the field of psychology. It is important to recognize that there are many varied opportunities to move forward.

References

- Barsky, R.B., Juster, F.T., Kimball, M.S. & Shapiro, M.D. 1997. Preference parameters and behavioral heterogeneity: An experimental approach in the Health and Retirement Study. *Quarterly Journal of Economics* 112(2) May: 537-579.
- Carstensen, L.L. 1993. Motivation for social contact across the life span: A theory of socioemotional selectivity. In J.E. Jacobs (Ed.) *Nebraska symposium on motivation: 1992, Developmental Perspectives on Motivation*, vol. 40 (pop. 209-254). Lincoln: University of Nebraska Press.
- Carstensen, L.L., Isaacowitz, D. & Charles, S.T. 1999. Taking time seriously: A theory of socioemotional selectivity. *American Psychologist*, 54, 165-181.
- Charles, S.T., Mather, M.M. & Carstensen, L.L. 2003. Aging and emotional memory: The forgettable nature of negative images for older adults. *Journal of Experimental Psychology: General*, 132(2): 310-324.
- Damasio, A.R. 1994. *Descartes' error: Emotion, reason, and the human brain*. New York: G.P. Putnam.
- Dominitz, J. 2001. Estimation of income expectations models using expectations with realizations data. *Journal of Econometrics* 102: 165-195.
- Dominitz, J. 1998. Earnings expectations, revisions, and realizations. *Review of Economics and Statistics*, August.
- Dominitz, J. & Manski, C. 2004. How should we measure consumer confidence? *Journal of Economic Perspectives* 18 (2): 51-66.
- Dominitz, J. & Manski, C. 2003. How should we measure consumer confidence (sentiment)? Evidence from the Michigan Survey of Consumers. *NBER Working Paper* No. W9926.
- Dominitz, J. & Manski, C. 1999. The several cultures of research on subjective expectations. In J.P. Smith and R.J. Willis (Eds.) *Wealth, Work, and Health: Innovations in Measurement in the Social Sciences*, Ann Arbor, MI: University of Michigan Press.
- Dominitz, J. & Manski, C. 1997. Using expectations data to study subjective income expectations. *Journal of the American Statistical Association*, 92: 855-867.
- Dominitz, J., Manski, C. & Fischhoff, B. 2001. Who are youth at-risk?: Expectations evidence in the NLSY-97. In R.T. Michael (Ed.) *Social Awakenings: Adolescents' Behavior as Adulthood Approaches*. Russell Sage Publications.

Dominitz, J., Manski, C. & Heinz, J., 2003 Will Social Security be there for you?: How Americans perceive their benefits. National Bureau of Economic Research Working Paper 9798.

Epstein, S. 1994. Integration of the cognitive and the psychodynamic unconscious. *American Psychologist* 49: 709-724.

Fischhoff, B., Parker, A., Bruine de Bruin, W., Downs, J., Palmgren, C., Dawes, R & Manski, C.2000. Teen expectations for significant life events. *Public Opinion Quarterly* 64: 189-205.

Fung, H.& Carstensen, L.L. 2003. Sending memorable messages to the old: Age differences in preferences and memory for advertisements. *Journal of Personality and Social Psychology* 85(1): 163-178.

Guiso, L., Jappelli, T. & Terlizzese, D. 1992. Earnings uncertainty, precautionary saving. Discussion paper, Centre for Economic Policy Research, United Kingdom. www.cepr.org/pubs/dps/DP699.asp

Guiso, L. Jappelli, T. & Pistaferri, L. 2002. An empirical analysis of earnings and employment risk. *Journal of Business and Economics Statistics* 20:241-53.

Hess, T.M., Pullen, S.M., & McGee, K.A. 1996. The acquisition of prototype-based information about social groups in adulthood. *Psychology and Aging* 11: 179-190.

Hibbard, J. & Peters, E. 2003. Supporting informed consumer health care choices: Data presentation approaches that facilitate the use of information in choice. *Annual Review of Public Health*, 24: 413-433.

Hsee, C.K. & Rottenstreich, Y. 2004. Music, pandas, and muggers: On the affective psychology of value. *Journal of Experimental Psychology: General* 133 (1):23-30.

Hurd, 1999, Experiments on anchoring and acquiescence bias, *Journal of Risk and Uncertainty*.

Hurd, M. & McGarry, K., 1995. Evaluation of the subjective probabilities of survival in the Health and Retirement Study." *Journal of Human Resources* 30: S268-S292.

Hurd, M. & McGarry, K., 2002. The predictive validity of subjective probabilities of survival. *The Economic Journal* 112: 966-985.

Hurd, M., Smith, J. & Zissimopoulos, J.M. Forthcoming. The effects of subjective survival on retirement and Social Security claiming. *Journal of Applied Econometrics*.

Jacowitz, K.E. & Kahneman, D.1995. Measures of anchoring in estimation tasks. *Personality and Social Psychology Bulletin* 21: 1161-66.

Kahneman, D. & Tversky, A. 1974. Judgement under uncertainty: Heuristics and biases. *Science* 185: 1124-31.

- Kahneman, D. & Tversky, A. 1979. Prospect theory: An analysis of decision under risk. *Econometrica* 47: 263-292.
- Kennedy, Q., Mather, M., & Carstensen, L.L. 2004 The role of motivation in the age-related positive bias in autobiographical memory. *Psychological Science* 15(3): 208-214.
- Locander, W.B., & Burton, J.P. 1976. The effect of question form on gathering income data by telephone. *Journal of Marketing Research* 13: 189-192.
- Lochner, L. 2003. Education, work, and crime: A human capital approach. Working Paper.
- Loewenstein, G.F., Weber, U., Hsee, C.K. & Welch, N. 2001. Risk as feelings. *Psychological Bulletin*, 2001. 127(2): 267-286.
- Manski, C.F. 2004. Measuring Expectations, *Econometrica* 72 (5):1329-1376.
- Manski, C. & Straub, J., 2000. Worker perceptions of job insecurity in the mid-1990s: Evidence from the Survey of Economic Expectations. *Journal of Human Resources* 35: 447-479.
- Mather, M., Canli, T., English, T., Whitfield, S., Wais, P., Ochsner, K., Gabrieli, J. & Carstensen, L. 2004. Amygdala responses to emotionally valenced stimuli in older and younger adults. *Psychological Science* 15: 259-263.
- Peters, E. & Slovic, P. 1996. The role of affect and worldviews as orienting dispositions in the perception and acceptance of nuclear power. *Journal of Applied Social Psychology*, 26(16): 1427-1453.
- Peters, E. & Slovic, P. 2000. The springs of action: Affective and analytical information processing in choice. *Personality and Social Psychology Bulletin*, 26(12): 1465-1475.
- Peters, E., Vastfjall, D., & Starmer, C. In review. Feeling your way to the right price.
- Rottenstreich, Y. & Hsee, C.K. 2001. Money, kisses, and electric shocks: On the affective psychology of risk. *Psychological Science* 12: 185-190.
- Tourangeau, R., Rips, L.J., Rasinski, K. 2000. *The Psychology of Survey Response*. New York: Cambridge University Press.
- Van der Klaauw, W. & Wolpin, K., 2002. Social Security, pensions, and the savings and retirement behavior of households. Department of Economics, University of North Carolina at Chapel Hill.
- Walker, J.R. 2001. Adolescents' expectations regarding birth outcomes: A comparison of the NLSY79 and NLSY97 cohorts. In R.T. Michael, ed. *Social Awakening: Adolescent Behavior as Adulthood Approaches*. New York, NY: Russell Sage Foundation (pp. 201-229).

Wennberg, J. & Cooper, M. 1999. *The Dartmouth Atlas of Health Care*. Chicago, IL: American Hospital Publishing, Inc.

Zajonc, Robert ,1997.

APPENDIX A

National Institute on Aging Workshop on Decision Making and Aging

Recommended Readings

Chuck Manski Recommends:

Manski, C. (2004) Measuring expectations, *Econometrica*.

Dominitz, J., Manski, C. & Heinz, J. (2003) Will Social Security be there for you? How Americans perceive their benefits (Working Paper).

Dan McFadden Recommends:

Hurd, M., McFadden, D., et al. (1998). Consumption and savings balances of the elderly: Experimental evidence on survey response bias, in D. Wise (ed.), *Frontiers in the Economics of Aging*, pp. 353 -387, Chicago, IL: University of Chicago Press.

M. Hurd, McFadden, D. & Gan, L. (1998). Subjective survival curves and life cycle behavior, in D. Wise (ed.), *Inquiries in the Economics of Aging*, pp. 259-305, Chicago, IL: University of Chicago Press.

Morikawa, T., Ben-Akiva, M. & McFadden, D. (2002). Discrete choice models incorporating revealed preferences and psychometric data, *Econometric Models in Marketing*, 16: 27-53, Oxford: Elsevier Science.

McFadden, D. (2002). Economic choices (Nobel Lecture), December. *American Economic Review*, 91 (3): 351-378, June 2001 and in *Les Prix Nobel*, 2001
[http://emlab.berkeley.edu/users/mcfadden/charterday01/charterday_final.pdf]

McFadden, D. (2004). Survey Design and Analysis (Gorman Lecture), June.
[http://emlab.berkeley.edu/users/mcfadden/charterday01/charterday_final.pdf]

McFadden, D., Schwarz, N. & Winter, J. (2004). Response behavior in surveys of the elderly: Experimental evidence from Internet surveys (conference presentation), June.

Green, D., Jacowitz, K., Kahneman, D. & McFadden, D. (1998). Referendum contingent

valuation, anchoring, and willingness to pay for public goods, *Resource and Energy Economics*, 20: 85-116.

Ellen Peters Recommends:

Hibbard, J. & Peters, E. (2003). Supporting informed consumer health care choices: Data presentation approaches that facilitate the use of information in choice. *Annual Review of Public Health*, 24, 413-433.

Kahneman, D. (2003). "A perspective on judgment and choice: Mapping bounded rationality." *American Psychologist* 58(9): 697-720.

Payne, J., Bettman, J., et al. (1999). "Measuring constructed preferences; Towards a building code." *Journal of Risk and Uncertainty* 19(1): 243-270.

Peters, E., Hess, T., Auman, C., Vastfjall, D. (in preparation). How do older adults decide?: Age differences in the impact of affect and deliberation.

Slovic, P., Finucane, M. L., Peters, E., & MacGregor, D.G. (2002). The affect heuristic. In T. Gilovich, D. Griffin, and D. Kahneman (Eds.), *Heuristics and Biases: The Psychology of Intuitive Judgment* (vol. 2) (pp. 397-420). New York: Cambridge University Press.

Peter Ubel Recommends:

Loewenstein, G.F., et al., Risk as feelings. *Psychological Bulletin*, 2001. 127(2): 267-286.

Ubel, P.A., Is information always a good thing? Helping patients make "good" decisions. *Medical Care*, 2002. 40(9 Suppl): 39-44.

Windschitl, P.D., R. Martin, and A.R. Flugstad, Context and the interpretation of likelihood information: The role of intergroup comparisons on perceived vulnerability. *Journal of Personality and Social Psychology*, 2002. 82(5): 742-755.

National Institute on Aging
Workshop on Decision Making and Aging
Gateway Building 5th Floor Conference Room
7201 Wisconsin Avenue
Bethesda, Maryland

July 14-15, 2004

Agenda
(Rev. 6-10-04)

July 14 (Wednesday)

7:00 PM **Group Dinner**
La Miche, 7905 Norfolk Avenue, Bethesda, Maryland (301-986-0707)

July 15 (Thursday)

8:00 AM **Coffee and Pastries**

8:30 AM **Welcome and Introductory Remarks**
Jeff Elias, National Institute on Aging

8:45 AM **Invited Perspectives**
Each speaker will have 15 minutes to present, guided by the following questions:
What is essential to know from the past to proceed?
Where are we now?
Where do we need to go?
What are the paths that can be followed to meet desired goals?

10:00 AM **Break**

10:30 AM **Invited Perspectives – Continued**

12:00 PM **Lunch**

1:00 PM **Group Discussion**

3:00 PM **Common Themes and Future Directions**

3:30 PM **Wrap Up**
Jeff Elias, National Institute on Aging

4:00 PM **Adjourn**